

## Durham Research Online

---

### Deposited in DRO:

27 August 2008

### Version of attached file:

Accepted Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Matthews, P. C. and Chesters, P. E. (2006) 'Implementing the information pump using accessible technology.', *Journal of engineering design*, 17 (6). pp. 563-585.

### Further information on publisher's website:

<http://dx.doi.org/10.1080/09544820600646629>

### Publisher's copyright statement:

### Additional information:

---

### Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

---

# Implementing the Information Pump using Accessible Technology

Peter C Matthews\* and Petrina E Chesters

School of Engineering, Durham University

Durham DH1 3LE, UK

April 2005

Contact address:

School of Engineering

Science Laboratories

South Road

Durham DH1 3LE, UK

Phone: +44 191 334 2538

Fax: +44 191 334 2377

email: [p.c.matthews@durham.ac.uk](mailto:p.c.matthews@durham.ac.uk) and [p.e.chesters@durham.ac.uk](mailto:p.e.chesters@durham.ac.uk)

---

\*corresponding author

## Abstract

The Information Pump (IP) is a method for extracting high quality subjective product information from a small group of subjects. The method is based around a game environment where the subjects are awarded points for the information they supply. This game environment provides incentives and motivation for the subjects to continue to provide high quality information throughout the game thus avoiding the fatigue issues that arise with other information elicitation methods. This paper develops the original IP method into one that can be implemented with commonly available office tools (paper forms and basic computing resources). This accessible version of the IP is implemented and tested through four separate games. The game outcomes are analysed both from the game player and product developer perspectives. These results indicate a sustained subject interest throughout the process and a set of high quality and varied product evaluation statements.

**Keywords:** Information elicitation; focus groups; non-technical product evaluation; subjective evaluation; game theory.

## 1 Introduction

The need for high quality, non-technical, product evaluation has been acknowledged for a long time (Griffin and Hauser; 1993; Yalch and Brunel; 1996; Lee et al.; 2005). These non-technical observations cover areas such as aesthetics, usability, desirability, fashion, etc., and are difficult to measure without recourse to the intended product market audience (Siu; 2005). A number of methods have been developed to facilitate this process, for example product focus groups, survey questionnaires and/or prototype evaluation (Spiggle; 1994; Takala; 2005). These methods tend to be costly and run the risk of boring the subject, resulting in poor quality information.

The Information Pump (IP) is a methodology that aims to counter the problems arising from traditional subjective product data collection Prelec (2001, 2004). The IP is a game theory based method that aims to maximise information extracted from a panel of subjects, while maintaining their interest in the process through a continuous panelist scoring method. The challenge with implementing the IP arises from the difficulty in executing the ‘game’. In its original format, there is an assumption that the game is played with each player using their own computer to interact with the game. While this in theory allows information and game scores to flow in a controlled manner between the players, it actually provides a major barrier to the wider adoption of the IP method due to the complexities involved in implementing and using such a system. This barrier is two-fold: it is costly and complex, and it is not an intuitive manner for the involved game players.

The IP method has been described as a ‘virtual focus group’, but it includes significant differences to set it apart from more conventional methods. It was developed in response to the growing concerns amongst academics and product designers that existing evaluation techniques available were not conducive to extracting usable information from the customer. It was believed that the lack of participant incentives resulted in low motivation levels of the participants involved in the evaluation process. Specifically, Resnick and Varian (1996) stated that ‘future systems will likely need to offer some incentive for the provision of recommendations’ in order to encourage respondents to make evaluations and impart their information.

The initial applications of the IP method have focused on the visual aesthetics of concept cars (Prelec; 2001; Dahan and Hauser; 2002; Matthews et al.; 2005). In particular, these initial tests concentrated on the function of the scoring system and incentive structure within the game. Four sets of participants evaluated products and other assorted visual stimuli using a paper-based protocol. In order to monitor the effect of the scoring system and incentive structure, one group evaluated the products using the complete version of

the IP whilst the second group used an interface that was identical in every respect apart from the exclusion of interactive scoring. In the original IP report, Prelec (2001) provides no detailed results to indicate the outcome or success of these preliminary investigations. However, according to Dahan and Hauser (2002), there were significant indications that the respondents using the absolute version of the IP provided statements and comments that were deemed by independent judges to be considerably more creative than those of their counterparts in the second group.

The core objective of this research was to develop and test a low cost version of the IP method. This used the game theory approach to maintain interest among players and maximise information extraction, but removed the need for each player to have their own computer interface to the game. This required replacing both the inter-player communication method and the score keeping/reporting. Finally, the results from the game process and player feedback were analysed to provide further recommendations for the implementation and limitations of the IP methodology.

## **2 The Need for Information**

Accurate and effective market research is very important to modern businesses in order to meet, and preferably exceed, customer expectations. For market research to be effective, several factors must be taken into consideration. These include the market at which the product is targeted, the method employed to gather information, and perhaps most importantly, the quality of the information gathered. Traditional methods of research include focus groups and surveys. Though there are benefits to each of these methods of gathering information, there are also flaws.

Lengthy surveys can lead to ‘participation fatigue’ i.e. boredom in participants. This symptom, along with difficult survey questions, leads to non-response and eventual attri-

tion from participants (Lee et al.; 2004). Providing an incentive, perhaps in the form of a prize or financial incentive, can reduce this participation fatigue, thus providing more information (Göriz; 2004). However, in cases where participation is not sufficiently supervised, consumers may complete the survey several times, or rush through the questions, in order to maximise their chances of gaining the incentive (Vehovar and Lozar; 1998). Similarly, incentives may be used to overcome low levels of participation in focus groups, but this does not guarantee participants with any real interest in the focus group topic (Morgan; 1998).

In both these instances, incentives only serve to improve the *quantity* of information gathered, not the *quality* of that information. The researcher assumes that participants will act in ‘good-faith’, delivering responses that they believe to be true.

## 2.1 Marketing Strategy

The most common reason for product failure is inadequate market analysis (Hollins and Pugh; 1990), leading to a poor marketing strategy. One of the most frequently cited examples of a product that failed due to poor market analysis is the Sinclair C5 (Bayliss; 1985; Hollins and Pugh; 1990; Dibb et al.; 2001). This was a small, one passenger, electric vehicle designed for urban transport. However, the vehicle failed to generate significant sales due to little interest in the product. If a product is to succeed, it is essential that the marketer find the correct marketing strategy for that product. The following stages are identified as being integral to the development of a successful marketing strategy (Zikmund; 2002):

1. Identify and evaluate opportunities;
2. Analyse and select the market segment, that is the section of the population to whom the product will be directed;

3. Plan and implement a marketing mix that will satisfy customer needs; and
4. Analyse marketing performance.

New opportunities can be identified in several ways. For example, a company may develop a new technology that has the potential to improve existing products (Hollins and Pugh; 1990). These are known as ‘Technology Push’ products. Alternatively, a new opportunity may come from the identification of an existing customer need. Products developed in this way are ‘Market Pull’ products. Regardless of how the product opportunity is identified, it must be thoroughly evaluated to ensure the product encompasses the customer need it is designed to fulfil. This can be achieved through marketing research, as described in Section 2.2.

The second stage of developing a marketing strategy is to analyse and select the market segment. In some cases, a firm may decide to direct the product at the whole market, focusing on what is common to every buyer as opposed to what sets them apart. This is known as an ‘Undifferentiated Strategy’ (Kotler and Armstrong; 1997). However, it is difficult to design a marketing mix that will appeal to such a wide range of potential buyers. Therefore, many marketers choose to implement a ‘Differentiated Strategy’ (Kotler and Armstrong; 1997). Specifically, a ‘Concentration Strategy’ may be chosen, aiming the product at just one market segment. Alternatively, a ‘Multi-Segment Strategy’, where the product is directed towards two or more segments, may be considered more appropriate (Dibb et al.; 2001). In order to do this, the marketer may choose to segment the market in any one of a number of ways. Examples include segmenting markets by age, sex, geographical location, or income (Kotler and Armstrong; 1997).

For the product to be successful within the chosen market segment, a marketing mix must be developed. The marketing mix is made up of four main components, known as the ‘Four Ps of Marketing’ (Bayliss; 1985). These are as follows:

**Product:** its features, advantages and disadvantages, expected life;

**Place:** position in the market, who will buy the product, how the product will be distributed;

**Promotion:** how the product will be made appealing to the customer, including packaging; and

**Price:** how the product's price compares with competing products.

When conducting marketing research, the marketer aims to elicit the information required to develop an attractive marketing mix that will appeal to a specified market segment. If a Concentration Strategy is implemented, it is necessary to develop a single marketing mix to suit the chosen segment. However, in Multi-Segment Strategies, different marketing mixes may be developed to suit different segments (Dibb et al.; 2001). For example, Proctor & Gamble markets nine different brands of laundry detergent, each with a different marketing mix aimed towards a different target market segment (Zikmund; 2002).

Finally, the marketer must analyse marketing performance. This may take the form of measuring actual sales against predictions, or may include further marketing research to establish whether the product meets customer expectations (Zikmund; 2002). If the marketing performance is poor, the marketer may adjust the marketing strategy in order to improve the profile of the product.

The Information Pump has potential to impact the first and third stages of developing a marketing strategy for a product. During the first stage, it is crucial to elicit relevant, high quality information from the customer to ascertain the needs that the product must fulfil. The IP would be executed at this stage using the concept design of a product, or other products in the same product category. The information elicited at this stage also



has applications when constructing the marketing mix during the third stage of strategy development. The IP will be particularly useful at this stage when the information elicited can be directly linked to the four Ps of marketing.

## **2.2 Marketing Research Methods**

When marketing a product, the marketer must be aware of all factors that influence a potential consumer's need or desire for that product. In order to gain the best possible view of consumer needs, marketers carry out marketing research. Marketing research can be conducted in many different ways, and used at different stages of the design process. During the planning and specification design phases, exploratory forms of marketing research are used to discover the basic needs and requirements that customers may have for a new product. Techniques include surveys or questionnaires, focus groups, depth interviews and thematic apperception tests.

Surveys are an inexpensive and simple way to gain information from a wide range of participants, particularly in the planning and clarifying design phase as an exploratory technique. However, the disadvantages are also numerous. Lengthy questionnaires can lead to 'participation fatigue' or boredom in participants. This, combined with difficult survey questions, leads to non-response and eventual attrition from participants (Lee et al.; 2004). Providing incentives can reduce this participation fatigue, allowing the marketer to gain more information from the survey (Göritz; 2004). However, in cases where participation is not sufficiently supervised, such as online surveys, consumers may complete the survey several times, or rush through the questions in order to maximise their chances of gaining the incentives (Vehovar and Lozar; 1998).

Focus groups and depth interviews allow a more supervised approach to data collection, and can be used either in the exploratory stages of research or during the conceptual design phase. Zikmund (2002) states that focus groups are fast, easy to execute, and inexpensive.

Morgan (1998), however, argues that time, effort, and an experienced moderator are required to gain useful information, along with time for post-research analysis. Experienced moderators are also a necessity when conducting depth interviews in order to gain relevant information from the interviewee. While incentives are common in enticing consumers to participate in both focus groups and depth interviews, this does not guarantee any interest in the topic of discussion (Morgan; 1998). In focus group sessions, there is also the risk that one participant may become dominant in the group, therefore discouraging other participants from expressing their true feelings (Zikmund; 2002).

A more unusual form of marketing research is the Thematic Apperception Test (Zikmund; 2002). During the test, consumers are asked to look at a series of pictures, and describe what the subjects in those pictures may do next. Themes are obtained from the ways in which participants respond. While this is a novel approach, designed to heighten participant interest in the subject matter, it is difficult to gauge the suitability of the pictures used until preliminary tests have been conducted (Zikmund; 2002).

Fundamentally, providing incentives in any one of the techniques mentioned above will only guarantee an improvement in the *quantity* of information elicited from the participant, not the *quality* of that information. While this can provide greater statistical significance to the data, the assumption remains that the subjects are acting in good faith, and assumptions are made about the quality of the information.

### **3 The Information Pump**

The Information Pump is a method to extract subjective information in the form of a controlled discussion group with incentives for high quality statements. A group of participants will view a product, and instead of simply voicing their own opinions, they will take turns to pose (‘encode’) statements that are relevant to the product. Each participant

then responds with their own binary view (agree or disagree) and attempts to guess the opinions of the other participants, using a scale of 1 (most disagree) to 9 (most agree). Depending on the accuracy of this prediction, the player gains or loses points. The scoring system introduces an element of competition into the game, with each player receiving a personal score.

A ‘dummy’ player, who cannot see the product visible to the other encoding players, is also present in the game. The dummy player is not required to make any statements, but must make a conscious effort to guess the responses of the other players. The more information the dummy learns throughout the game, the better they can guess and hence the fewer points the encoders can score. Scoring is designed to provide incentive to the encoders to vary their statements, which results in a richer information source in terms of product evaluation.

At the end of the game, all scores are accumulated, and the winner is the player who has provided the most qualitative and quantitative information during the game and has been the best judge of others. This player can be noted as a ‘good player’ and can be encouraged to return for future product research. The designer can now use the information generated during the game as product evaluation data.

### **3.1 Game Theory**

Game Theory is used widely in economics and other disciplines to describe situations where there is competition between inter-dependent participants (or players). Each player has a set of actions available to them, from which a set of strategies can be derived depending on the actions of the other players (Romp; 1997). For example, consider the simple game where one player owns a car and another wishes to purchase the car. The potential buyer provides an offer to the car owner. The car owner has two available actions: to either agree to sell the car for a given offered price or to reject the offer. In this game, a simple

strategy for the owner would be to accept the offer if it is greater than, say, £500, or to keep the car if the offer is less than that. In a static game, players make their decisions simultaneously without knowing what decisions the other players have made. In the case above, this would result in the car owner deciding to sell or keep the car *before* knowing how much the buyer will pay and hence must base the decision on what the owner *believes* the buyer will offer.

During each game modelled using game theory, there are payoffs or points scored. These payoffs are designed so that each player will strive to gain the highest payoff possible, but the payoff depends on the decisions of all players in the game. It is assumed that all players are rational in attempting to maximise their payoff, and that each player therefore expects all other players to be rational. A ‘zero sum’ game is where the payoffs of one player are made at a direct loss of another player’s payoff, and hence the sum of all the points gained at any point in the game is zero. A game where the sum of all players’ payoffs is not equal to zero is called a non-zero sum game.

The most common example of a non-zero sum game is the ‘prisoner’s dilemma’. In this game scenario, two men are being questioned by the police about a crime. They can either remain silent or defect by providing evidence incriminating the other prisoner. The police do not have sufficient evidence to convict either, so if neither man defects, both will be sentenced to three months. If only one man defects, he will go free while the other will be sentenced to nine months in jail. If both co-operate and confess, they will each be given a six-month prison sentence. Table 1 represents the pay-offs for both prisoners according to the actions they take. While the best option is for both to remain silent, this is an unstable solution: either prisoner would do better by not co-operating with the other prisoner. Hence, it is not a rational solution. The rational solution is where both prisoners confess: each could do no *worse* than if they chose to be silent.

In games where there are more than two players, or many more actions and strategies

Table 1: Payoff table for the Prisoner’s Dilemma: The values represent the time for the (row, column) prisoner’s time in jail. A prisoner ‘defects’ by talking to the police as opposed to remaining totally silent.

	Defect	Silent
Defect	(6,6)	(0,9)
Silent	(9,0)	(3,3)

are available to players, it may not be so simple for players to derive the most rational action to take. In situations such as this, players will take into account the risk they feel is associated with each available action, and choose a strategy accordingly (Yates; 1992). Some players will repeatedly choose the action they feel to be the least risky, and are known as being ‘risk averse’. Others will choose risky options most often in an attempt to gain the highest payoffs. These players are known as being ‘risk seeking’. The final group of players are known as being ‘risk neutral’. These players balance risk with payoff as well as possible at every opportunity.

When participating in the Information Pump, players are competing against each other to gain the highest payoffs. The IP can be modelled as a non-zero sum game. As such, a good judgement made by one player is not directly detrimental to another player. Likewise, a bad judgement made by one player is not directly beneficial to another. However, consistent good judgement throughout the game will lead to good payoffs, and a high score at the end of the game.

As the IP is a complex game, and cannot easily be solved as shown for the prisoner’s dilemma, players’ attitudes to risk will be demonstrated during the game. For example, the extreme response values of 1 and 9 offer players the opportunity for the greatest payoffs, but also pose the greatest risk of losing points. A risk-averse player may never use these response values, preferring to be more certain of a smaller payoff. A risk-seeking player, however, is likely to use the values 1 and 9 regularly.

### 3.2 Information Pump: Details

The Information Pump is a non-zero sum game (Prelec; 2001, 2004). The overall aim of the game for the product researcher is to generate an interesting and informative set of statements about a product, along with a perceived ‘truth’ value for each statement. This is achieved through a competitive game context where points can be scored for (1) interesting statements and (2) honest evaluation of those statements. The overall aim for the players is to accumulate as many points as possible.

The game has two types of players: ‘encoders’ who have knowledge of the product and ‘dummies’ who do not. The encoders take turns to use their knowledge of the product to create statements describing the product. These statements are broadcast to all players. The other encoders (acting as ‘decoders’) individually decide if the statement is true or false and make a forecast on how the remaining decoders will perceive the truth of the statement. The dummies only forecast how the encoders will vote.

The aim of the game is to create statements that are unambiguous for those with knowledge of the product but are ambiguous without that knowledge. This provides the challenge for the encoder: the statement must be recognised by the players with product knowledge (i.e. the decoders) but not by the dummies. All players have access to (or knowledge of) all previous statements. Thus, if an encoder makes a redundant statement, the dummies can use the previous statements to begin to gather information about the product and as a result be able to better forecast how the encoders will perceive the statements.

Points are scored according to how well players forecast the overall encoders’ perception of the truth value of a statement. The dummies’ score ( $s_{\text{dum}}$ ) is determined based on how well they forecast the encoders’ perception. Effectively, this rewards the dummies for learning about the product. In a ‘good’ game, the dummies’ score is expected to be zero at the end of the game. The decoders’ score ( $s_{\text{dec}}$ ) is computed similarly to the

Table 2: Payoff look-up table taken from Prelec (2001).

Forecast	1(F)	2	3	4	5	6	7	8	9(T)
$l_T$	-70	-40	-22	-10	0	8	15	20	25
$l_F$	25	20	15	8	0	-10	-22	-40	-70

dummies', but subtracting the dummies' average score to counter-balance the decoders' product knowledge. Thus, a positive score indicates that the decoder has performed better than the dummies, and vice versa for a negative score. Finally, the encoder's score ( $s_{\text{enc}}$ ) is computed based on the decoders' average score less the dummies' average score. Hence, the encoder is rewarded for creating statements that are clear to the decoders but not to the dummies.

Depending on what type of player  $j$  is (i.e. dummy, decoder or encoder), the scores from statement  $i$  for player  $j$  are computed explicitly as follows:

$$s_{\text{dum}}(i, j) = |T(i)|l_T(f_{ij}) + |F(i)|l_F(f_{ij}) \quad (1)$$

$$s_{\text{dec}}(i, j) = |T(i) \setminus j|l_T(f_{ij}) + |F(i) \setminus j|l_F(f_{ij}) - \frac{\sum_{k \in D} s_{\text{dum}}(i, k)}{|D|} \quad (2)$$

$$s_{\text{enc}}(i, j) = \sum_{k \in T(i) \setminus j} l_T(f_{ik}) + \sum_{k \in F(i) \setminus j} l_F(f_{ik}) - \frac{\sum_{k \in D} s_{\text{dum}}(i, k)}{|D|} \quad (3)$$

where  $T(i)$  and  $F(i)$  are the sets of encoder/decoder player indices who evaluated statement  $i$  as true and false, respectively;  $D$  is the set of dummy player indices;  $|X|$  is the number of elements in set  $X$ ;  $f_{ij}$  is player  $j$ 's forecast on statement  $i$ ; and  $l_T$  and  $l_F$  are the payoff values taken from a pay-off look-up table, for example Table 2.

The Information Pump rewards honest answers. The encoders maximise their score by posing statements that they believe they will be best able to predict how the other knowledgeable decoders will answer and that will reveal the least to the dummies. The

Table 3: Payoff look-up table taken from Robson (2004), removing the neutral strategy.

Forecast	1(F)	2	3	4	5	6	7	8	9(T)
$l_T$	-65	-35	-18	-5	0	7	12	19	30
$l_F$	30	19	12	7	0	-5	-18	-35	-65

decoders maximise their score by agreeing with what they believe is the common view of the other decoders. Where decoders believe there is ambiguity in the statement (i.e. they are unsure the rest of the decoders will vote as they feel), this is reflected by forecasting a more central position, a less risky position for the decoders. The dummies maximise their score by carefully gaining as much information from the previous statements as possible and using this to improve their forecasts throughout the game.

The dummies play an important role in the Information Pump. During the game, they provide a tangible challenge to the encoders to ensure that they do not reveal any information about the product. This ensures that encoders minimise making redundant statements about the product that the dummies could use to gain product knowledge. In a ‘good’ game where no knowledge is leaked to the dummy, the dummies expected score is zero: namely the dummies were guessing throughout the game. Where knowledge has been leaked out or redundant statements are being made, the dummies on average begin to score. The information from the dummies’ scores can also be used when post-processing the statements: where dummies begin to perform well provides an indication that the quality of the information being generated in the statements is beginning to wane.

The ultimate aim of the IP is to extract high quality subjective information about a product. The game format provides a means to this end. The direct output of the game is a series of statements and the associated responses for each statement. These statements provide the subjective perception of the product. Each statement can be roughly classified in three categories depending on the associated responses: clearly true, clearly false, and



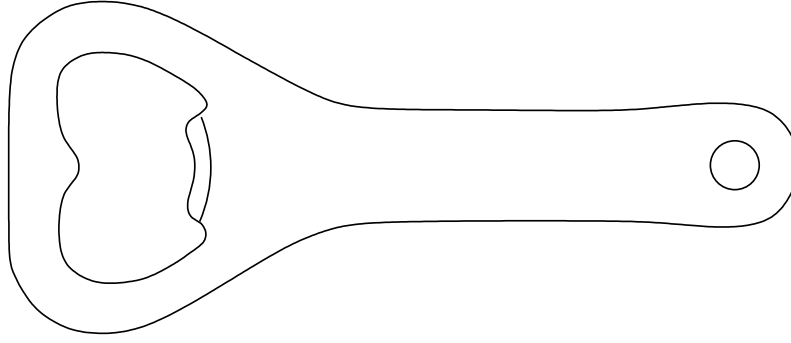


Figure 1: Schematic diagram of the ‘classic’ stamped bottle opener.

ambiguous. The two ‘clear’ categories represent the strong image that a product projects. The ambiguous statements represent areas where the projected image is ambiguous. Recall, the encoder posing the statement wished to maximise their score by providing unambiguous statements. However, for the ambiguous statements, the remaining players will have been divided on their views regarding the statement. This is useful information to the designer and/or marketer, as it indicates areas the product does not clearly project. Where this is felt to be important, the designer can invest more effort in providing a clear message.

### 3.3 Game trace

To illustrate the running of the Information Pump, a segment of a fictitious game will be described. This fictitious game will consider a simple bottle opener as the product. This is the ‘classic’ simple version of the stamped bottle opener, as illustrated in Figure 1. It is a thin bottle opener with few aesthetic properties. This product has no other function than opening capped bottles.

This game is played with six participants. Each of these participants will be allocated one role: one dummy player, one encoding player, and four decoding players. The encoding player can be swapped with one of the decoding players during the execution of the game. However, the dummy player must remain in that role throughout the game, but this is

Table 4: Player responses from the fictitious game. The dummy only reports his perception of the decoders’ and encoder’s average. The decoders and encoder report their perception of the statement and their confidence that this perception is shared between the panel.

Stmt	1–Dummy	2–Dec <sub>1</sub>	3–Dec <sub>2</sub>	4–Dec <sub>3</sub>	5–Dec <sub>4</sub>	6–Enc
1	5	T9	F3	T8	T7	T9
2	5	F3	F2	F3	F1	F2
3	5	F4	T5	F6	F3	T8
4	9	T8	T8	T7	T9	T8

not done in this game. The dummy must be seated in a manner that will not enable any stimulus or knowledge, visual or otherwise, to be obtained from the product under consideration. The players’ ‘moves’ from the game are in Table 4.

Once the dummy has been selected, the players seat themselves such that the encoding and decoding players have access to the product, but the dummy does not. This is achieved by seating the dummy behind a screen, so that the dummy can still participate verbally but not see the product. This is the basis of the scoring mechanism: encoder and decoder scores are maximised when statements are made that are clear to those who have knowledge of the product, but are totally without meaning for the dummy who has no knowledge of the product. A ‘good’ game is characterised by one where the dummy’s score is near zero.

Once all players are seated, the product is introduced to the encoder and decoders. For illustration purposes, a total of four product statements will be made by the encoder. For the first statement, the encoder decides to describe the main functional aspect of the product: *“The product serves a single purpose”*. For most of the decoders, this statement appears clearly true and they mostly respond as such with high confidence that their perception of this statement will be shared amongst the rest of the decoders. One decoder (2) believes this product could serve other purposes, and reports this statement as being false. The dummy has knowledge of the product, and hence must guess how the other players perceive this statement. As the dummy is making a total guess, this is reflected by the

central view (or low confidence position) the dummy has of the panel with respect to this statement. The statement is published so that all players, including the dummy, can refer to it later. This is useful for the dummy to review the history of statements to potentially gather product information and for the future encoders to ensure that they minimise information leakage through redundant statements.

The scores are computed using Equations 1–3. The relevant sets used by these equations for this statement are:  $T(1) = \{2, 4, 5, 6\}$ ,  $F(1) = \{3\}$ ,  $D = \{1\}$ , and the player confidences for this first statement are  $f_{1j} = (5, 9, 3, 8, 7, 9)$  (where  $j$  is the index within this array). The scores for each player are explicitly computed as follows, using the lookup table values taken from Table 2:

$$\begin{aligned}
s_{\text{dum}}(1, 1) &= |T(1)|l_T(f_{11}) + |F(1)|l_F(f_{11}) \\
&= 4 \times (0) + 1 \times (0) = \mathbf{0} \\
s_{\text{dec}}(1, 2) &= |T(1) \setminus 2|l_T(f_{12}) + |F(1) \setminus 2|l_F(f_{12}) - \frac{\sum_{k \in D} s_{\text{dum}}(1, k)}{|D|} \\
&= 3 \times (25) + 1 \times (-70) - 0 = \mathbf{5} \\
s_{\text{dec}}(1, 3) &= 4 \times (-22) + 0 \times (15) - 0 = \mathbf{-88} \\
s_{\text{dec}}(1, 4) &= 3 \times (20) + 1 \times (-40) - 0 = \mathbf{20} \\
s_{\text{dec}}(1, 5) &= 3 \times (15) + 1 \times (-22) - 0 = \mathbf{23} \\
s_{\text{enc}}(1, 6) &= \sum_{k \in T(1) \setminus 6} l_T(f_{1k}) + \sum_{k \in F(1) \setminus 6} l_F(f_{1k}) - \frac{\sum_{k \in D} s_{\text{dum}}(1, k)}{|D|} \\
&= (25 + 20 + 15) + (15) - 0 = \mathbf{75}
\end{aligned}$$

These scores demonstrate that where the dummy gains no knowledge, no scores will be accrued by the dummy. The decoders gain points for judging how the others will perceive the statement, and conversely lose points where they misjudge this. Finally, the encoder gains points for providing a clear statement that the decoders feel confident about judging

the panel's perception.

For the second statement, the encoder decides to describe the cost appearance of the product: *"The product is expensive"*. Again, this statement is clear to the decoders who all perceive this as being false and are confident that this view is shared with the other decoders. Again, the dummy has no means of determining the perception by others of this statement and must guess with no confidence. These scores are computed as follows (presented briefly):

$$\begin{aligned}
s_{\text{dum}}(2, 1) &= 0 \times 0 + 5 \times 0 = \mathbf{0} \\
s_{\text{dec}}(2, 2) &= 0 \times (-22) + 4 \times (15) - 0 = \mathbf{60} \\
s_{\text{dec}}(2, 3) &= 0 \times (-40) + 4 \times (20) - 0 = \mathbf{80} \\
s_{\text{dec}}(2, 4) &= 0 \times (-22) + 4 \times (15) - 0 = \mathbf{60} \\
s_{\text{dec}}(2, 5) &= 0 \times (-70) + 4 \times (25) - 0 = \mathbf{100} \\
s_{\text{enc}}(2, 6) &= (\emptyset) + (15 + 20 + 15 + 25) - 0 = \mathbf{75}
\end{aligned}$$

With the third statement, the encoder attempts to describe what sort of person would be attracted to this bottle opener. The encoder, believing that a boy scout would be attracted to the simplicity of the design, states: *"A boy scout would always have this product on his person"*. However, some decoders believe that the boy scout would be disappointed with the simplicity, preferring a multi-functional tool, and disagree with the encoder's view. Similarly, the decoders are not convinced that this is a clear statement and are unsure of how others will perceive this statement. As a result, they also score their perception of others' view as unsure. Again, the dummy can make little of this statement and must guess. The scores for this round are:

$$s_{\text{dum}}(3, 1) = 2 \times 0 + 3 \times 0 = \mathbf{0}$$

$$\begin{aligned}
s_{\text{dec}}(3, 2) &= 2 \times (-10) + 2 \times (8) - 0 = -4 \\
s_{\text{dec}}(3, 3) &= 1 \times (0) + 3 \times (0) - 0 = \mathbf{0} \\
s_{\text{dec}}(3, 4) &= 2 \times (8) + 2 \times (-10) - 0 = -4 \\
s_{\text{dec}}(3, 5) &= 2 \times (-22) + 2 \times (15) - 0 = -14 \\
s_{\text{enc}}(3, 6) &= (0) + (8 - 10 + 15) - 0 = \mathbf{13}
\end{aligned}$$

As can be seen from these scores, an ambiguous statement results in very low scores for both the decoders and encoders.

In the final statement, the encoder begins to struggle to identify suitable statements, and decides to describe the product's context. The encoder provides the statement: *"The product provides means for drinking"*. The decoders clearly and confidently agree with this statement. This time, the dummy uses previous statements to form a picture of the product. In view of this drink related statement, the first two statements (single purpose and cost) suggest the product could be a glass or some means of accessing a drink. The third statement (boy scout) biases the statement away from the glass, as the dummy assumes that boy scouts drink directly from bottles. As a result, the dummy confidently agrees with the statement. The scores for this statement are:

$$\begin{aligned}
s_{\text{dum}}(4, 1) &= 5 \times (25) + 0 \times (-70) = \mathbf{125} \\
s_{\text{dec}}(4, 2) &= 4 \times (20) + 0 \times (-40) - 125 = -45 \\
s_{\text{dec}}(4, 3) &= 4 \times (20) + 0 \times (-40) - 125 = -45 \\
s_{\text{dec}}(4, 4) &= 4 \times (15) + 0 \times (-22) - 125 = -65 \\
s_{\text{dec}}(4, 5) &= 4 \times (25) + 0 \times (-70) - 125 = -25 \\
s_{\text{enc}}(4, 6) &= (0) + (8 - 10 + 15) - 125 = -45
\end{aligned}$$

From this final statement, where the dummy has made a bold and correct estimate on the perception of the remainder of the panelists, it can be seen that the scores are significantly different. The dummy obtains a very high score for correctly estimating the panel's view. This in turn penalises the remainder of the panel for effectively not producing 'new' information.

## 4 Empirical Methodology

The overall aim of this research is to develop and test an improved version of the paper-based IP. This broad aim can be broken down into objectives as follows:

1. To devise a system to inform players of their scores throughout the game, with the intention of improving incentives;
2. To implement the IP method in two experiments using a paper-based method; and
3. To analyse the results obtained in both experiments, and suggest further improvements based on this analysis.

### 4.1 Hypotheses

The IP relies on participants being interested and engaged at all times. One factor that differentiates this method of data extraction from others is that the model offers the participants incentives in the form of a possible prize at the end of the process. As part of Prelec's web-based model, participants are able to track their scores and the scores of the other players throughout the game. However, in previous paper-based versions of the IP this has not been the case, with players waiting until the end of the game to discover which participant has won (Matthews et al.; 2005). Experiment 1 sought to investigate the impact of allowing players to track the scores of all players as the game progresses.

It was anticipated that being able to see a cumulative score as the game progressed would have had the following impact: players would be able to better assess their own strategies; players will remain engaged and focused for longer; and the competitive nature of the game will bear through.

## 4.2 Empirical Methods

At the core of the empirical method lies the aim of creating a simple and accessible platform for implementing the IP. Thus, the equipment required was limited to a computer with a spreadsheet application, a suitable room, and suitable furniture for the players and moderator (acting as an observer/researcher). Figure 2 illustrates how the furniture was arranged in the room. The room had capacity for five players, one of which would be the dummy. A screen was erected to visually shield the dummy from the encoders and the product under consideration. The encoders could also refer to a set of keywords or other concepts that could be used as a source of inspiration for generating statements. These keywords provided a means for the moderator to direct the nature of information extracted. Encoders were prevented from seeing other encoders' written responses, as this would provide a favourable bias towards the 'cheating' encoder's score, although it would not directly affect the quality of information generated through the game. Towards the 'front' of the room, the previous statements and scores were visible to all players. The moderator was free to move about the whole room as needed, however this was minimised whenever possible.

In addition to the room setup, during some games the players were provided with supplemental information. This was provided in printed format and covered potential strategies, payoff tables, and other material that might influence the nature of the game and as a result, the quality of the information generated during the game.

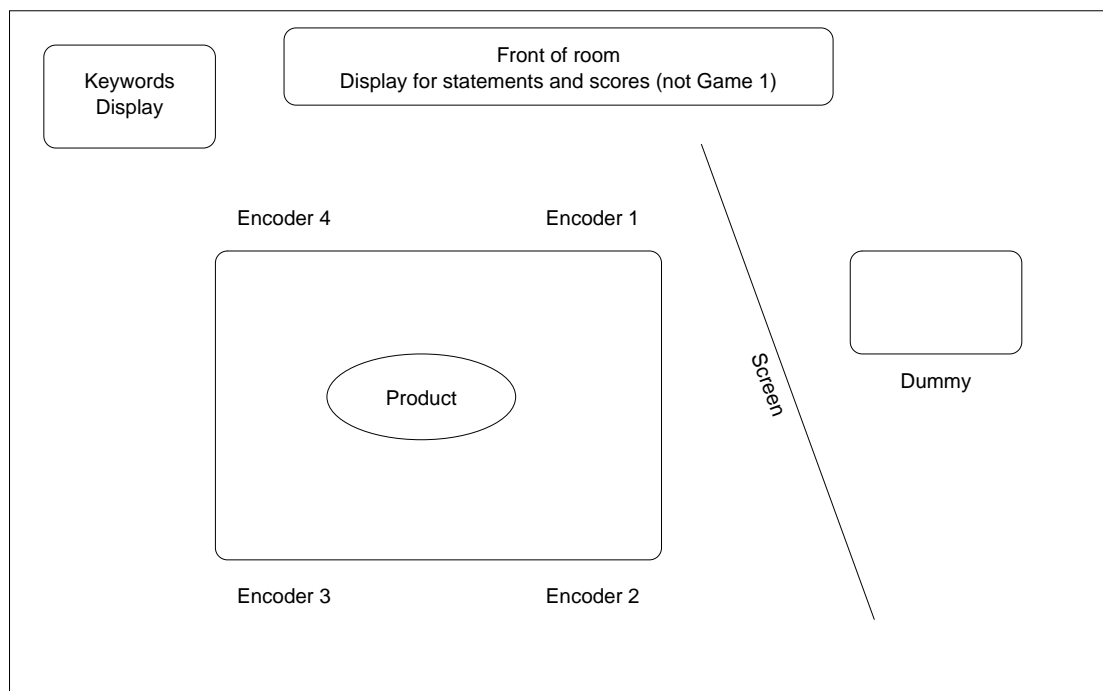


Figure 2: Room layout for the Information Pump experiments. Note that the statements and scores were not displayed for Game 1.



### 4.3 Measures and Data capture

There are two quantitative aspects to running the IP game: (1) scoring the players for each statement according to the responses given by each player, and (2) scoring the individual statements to provide a ‘quality of information’ measure.

The player scoring is primarily a device for maintaining the players’ interest in the information gathering process. By providing explicit scores throughout the game, a competitive element is brought to the proceedings. Ideally, the players have a good understanding that generating high quality information results in achieving higher scoring rounds. However, these player scores have little effect on how the information is used after the game.

The scoring of the individual statements generated during the game provides a measure of the overall subjective perception of each individual statement. Thus, the statement and its associated score provide feedback to the designer and marketer of some subjective perception of the product. These statements need to be analysed after the game to obtain the overall subjective view of the product under analysis.

As the aim of this work was to investigate and develop accessible methods for executing the IP, the primary data collection was done through paper forms. For each statement, the players fill out one form (see Figure 3 for the encoder/decoder form and Figure 4 for the dummy’s form). As each statement is made by the encoding player, it is given a number and recorded by the game moderator, and written down on the public notice board for all players to refer to. The encoder/decoder players (who have access to the product) are asked to provide (1) their opinion as to the truth of the statement and (2) how they *think* the other encoders/decoders will evaluate the statement. The dummy is asked only about how they believed the encoders/decoders perceived the statement.

Finally, after each experiment, a post-experiment questionnaire was filled in by all the players. This was used to evaluate the players’ perception of the game process, rather than to evaluate the products. Six aspects of the game were evaluated: enjoyment of the game;

Please enter the statement number

Are you the author of this statement?

Please indicate whether you believe the statement is TRUE or FALSE

Please indicate how you think the other players perceive this statement (tick one)

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

(FALSE) (TRUE)

Figure 3: Paper form for capturing encoder/decoder perceptions.

Please enter the statement number

Please indicate how you think the other players perceive this statement (tick one)

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

(FALSE) (TRUE)

Figure 4: Paper form for capturing dummy perceptions.

understanding of the game; understanding of scoring; preference for knowing scores; ease of filling out forms; and helpfulness of keywords. These were scored on a scale of 1–very low to 5–very high.

## 5 Empirical Work

The context of the empirical work was determined by the scope of the research project’s constraints. This project represented the research component in the final year of a Masters of Engineering degree course. As such, the research had limited access to products. The products that were considered were largely new market entrants. The Information Pump would normally be used at a more conceptual stage in the design process, however there was no access to products at this development stage. This had little effect on the empirical research, as most of the products used were recent market entrants and were not known by the players.

In each experiment, five players were taken from the student population. While efforts were made to ensure the players had not met before, this was not completely possible. The aim was to minimise uncontrolled communication between players that might occur due to some players being aware of other players’ habits.

Before each experiment started, the players were supplied with a written overview of the Information Pump detailing the aims, objectives, rules of the game and possible game strategies. Any queries the players had were addressed after the players had reviewed the game documentation, but before the game started.

Each experiment consisted of two games, each game considering a different product. A dummy was selected at random from the five players, and this player served the dummy role for just one game. On selection, the dummy was seated behind the screen and the remainder of the players (who were now encoders) took their seats. Once all seated, the

product was revealed to the encoders. The level of interaction the encoders had with the product was determined by the nature of the product.

## 5.1 Experiment 1

The first experiment was designed to test two main areas: what effect did incentives have on player participation during the game and how effective was the scoring system. The IP is designed to maintain player interest throughout the game, and this experiment was designed to provide evidence of this. Previous work noted that the lack of directly available scores did suggest a waning interest in the game (Matthews et al.; 2005).

The first experiment consisted of two games, each investigating a different product. The two games were to be run slightly differently in order to test the following hypotheses:

1. Players will be better able to assess the success of their own strategy and that of others, and alter their tactics accordingly;
2. Players will be able to learn tactics other than those suggested as the game progresses;
3. Players will remain engaged and focused for longer periods of time during the game due to the combination of an overall incentive to win, and smaller incentives to gain a good score for each statement; and
4. The competitive nature of the Information Pump will be emphasised, as players will be aware of their rank position.

It is also possible that those players that struggle to understand the game may become disinterested more quickly, as they may fail to accumulate the higher scores of those players who do understand. The hypotheses were measured by playing one game of the Information Pump with no scores displayed to the players, followed by a second game where the scores

were displayed. At the end of the game, players were asked to complete a questionnaire comparing the features of both games.

#### **5.1.1 Game 1**

The product evaluated in Game 1 was ‘Brunchettas’, a novel food product where the user applies a cheese spread to a flavoured biscuit using a supplied spreader. The product is marketed as a compact, convenient and nutritious snack. The manufacturer describes the product as being an ‘anytime delicatessen’, primarily targeted for work-time consumption.

Before the game started, a dummy was selected at random from the five players. This dummy was placed behind a screen, unable to communicate at any level with the encoders. Communication between encoders was strictly monitored, and it was made clear that any queries must be made via the game moderator. The product was revealed once all players had taken their seats. During the first half of the game, the encoders were only allowed to look at the packaging. Approximately half way through the game, the players were allowed to open the package and sample the product.

During the game, the scores were computed using a spreadsheet model. These scores were only revealed at the end of the game. Players were also not provided with any information as to how their scores were computed.

#### **5.1.2 Game 2**

The product evaluated in Game 2 was ‘Babycham’, which has a long history in the British alcoholic beverage market. Babycham was the first ‘branded’ alcoholic drink, and entered the British market in the 1950s. Its target was young women who found the light, sweet, and sparkling drink appealing compared with the alternatives then available. Babycham’s success continued throughout the 1960s and 1970s, but its popularity declined in the 1980s. However, towards the end of the 1990s, a trend of ‘retro’ style products emerged, and

sales of Babycham increased. A new range of products was launched for the millennium, including bottles with popping corks to emulate champagne. The producers of Babycham believe that the brand has come full circle, and is now a drink that a whole new generation will choose.

Game 2 used the same players as were used for Game 1. Prior to the start of Game 2, all players were given supplementary information regarding the scoring system. This information included both a written explanation of the system and numerical examples to illustrate the explanation. Once all players had read the information, any questions were answered by the moderator so that all players were satisfied they had a clear understanding of how the scores would be computed.

Again, a dummy was chosen at random before the game commenced and asked to sit behind the screen. Similarly to Game 1, mid-way through the game the encoders were allowed to open and sample the product. However, in Game 2 the player scores were publicly displayed and updated after each statement was posed. In addition, the Robson (2004) pay-off table was used (see Table 3), which inhibits selecting neutral positions.

## 5.2 Experiment 2

Experiment 2 consisted of two Information Pump games, Game 3 and Game 4. These games incorporated a number of modifications, including obtaining the encoders' opinion on statement relevance. This had no effect on the scoring system, but provided additional information for the post-game statement analysis.

Five players participated in this experiment, all players taking part in both games. Before the first game, the players were provided with full instructions on playing the game including strategies and scoring payoffs. After the experiment, the players were asked to fill out a questionnaire regarding their opinions of the Information Pump game.

Experiment 1 provided useful product information, however it was not clear how rel-

evant this information was perceived to be and it was difficult to fit the statements into predetermined categories. It was also noted that the players did not use the keywords provided, suggesting that the Experiment 1 games lacked guidance. Thus, the main objectives of Experiment 2 were to investigate:

1. The relevance of statements made in the Information Pump games to both the products being discussed and the information required by marketers; and
2. Methods other than the use of keywords to guide players in providing information useful in the marketing of the product.

In order to quantify individual statement relevance to the product, encoders were additionally asked to score their perception of the relevance of each statement on a scale of 1–very low to 5–very high. To provide greater guidance to the players, the keywords displayed were taken from the ‘Four Ps of Marketing’, as described in Section 2.1 (Bayliss; 1985). In Game 3, the encoders were required to make statements under specified categories. In Game 4 the encoder posing the statement would note down which category the statement belonged to, rather than be required to make statements within a certain category.

It was hypothesised that:

1. Players will be more likely to provide relevant information relating to the product in the categories defined during both games;
2. Players will find the forced method of Game 3 restrictive to creativity in statement composition; and
3. Players will find the less structured format method of Game 4 more appealing, leading to enhanced creativity in statement composition.

These hypotheses were measured using the relevance scale on the encoder/decoder response sheets, and player responses to the post-experiment questionnaire.

### **5.2.1 Game 3**

The product used in Game 3 was ‘Kombucha’, a non-alcoholic drink that is part of the Carpe Diem brand. The manufacturers describe Carpe Diem as both a brand and a ‘philosophy’. Its beverages are designed to be part of a lifestyle choice. The brand website offers advice to consumers on their ideal ‘home spa’, with emphasis being placed on Carpe Diem drinks as part of that relaxing environment.

Kombucha is one of three products in the Carpe Diem range, all of which have some association with spiritual well being. Kombucha is described as a ‘mystical drink, long a part of the ancient philosophy of Zen’. It is designed to appeal to those who envisage themselves as being spiritually balanced, particularly women who feature strongly on the brand website.

At the start of Game 3, a dummy player was chosen at random and isolated from the rest of the players. All players were then allowed to clarify any questions they had about the game procedure with the moderator. A board was displayed to the encoders on which was written a list of four categories relating to the four Ps of marketing. The first two rounds of the game were devoted to the first category and so on, requiring each encoder to make two statements in each category. Players responded to the statements using the response forms provided.

Half way through the game, the encoding players were allowed to open the product in order to see, smell, and taste it.



### 5.2.2 Game 4

The product used in Game 4 was the ‘Toilet Duck Fresh Brush’, a new innovation from the SC Johnson home cleaning products company. It is designed to compete with the traditional toilet brush and other similar products from competing companies that are also new to the market. The concept is simple: the Toilet Duck Fresh Brush uses cleaning pads that can be disposed of after a single use. The appeal of this is that it is perceived to be more hygienic than traditional toilet brushes. The target market for the Toilet Duck Fresh Brush is the hygiene-conscious homeowner.

At the beginning of Game 4 a different dummy player from the Game 3 dummy was selected at random. This player was isolated for the duration of the game. During Game 4, the same four marketing categories were displayed as in Game 3. However, the moderator did not dictate to the players the number of statements to be made under each category heading or the order in which categories should be addressed. As each encoder made a statement, they indicated which category they felt the statement came under.

The players responded to each statement using the response forms provided. Again, a relevance measure was included, which players used to indicate how relevant they felt the statement was to the product and the category in which it was placed. At the half-way point in the game, the encoders were able to open the product and assemble the parts in order to see how the product looked and worked.

## 6 Interpretation of Results

The results need to be analysed from two perspectives: the product related information elicited and the participants’ opinions of the IP game process. The two experiments collected data on product statements and associated player perception; and post-experiment questionnaires. The second experiment also collected data on the perceived relevance of

each statement.

## 6.1 Product perception analysis

The bulk of the data generated came from the product statements and the associated player perceptions. For the first experiment, each data point consists of three elements: the statement, the player perceptions (coupled with their confidence), and the total score for the statement. In the second experiment, a perceived statement relevance was included in each data point.

Game 1 generated a total of 32 data points. These primarily consisted of statements which the encoders were in agreement on (as indicated by high scores), but also included a number of statements where the players did not agree upon although the disagreement was weak. The high scoring statements provide information about how the panel, and by extension the general public, perceive the product. Table 5 contains the ‘Placement’ relevant statements made by the encoders for Game 1. The high scoring statements provide an insight into how the product is publicly perceived: namely that it is a new snack food which is not appropriate for consuming in a rush. The low scoring statements highlight areas that are less clear: whether the public would purchase the product and if the product would be bought for children. From these low scoring statements, it is apparent that the product requires to project a clearer image with respect to the product placement perception.

Game 2 generated a total of 24 data points. Table 6 contains the three statements relevant to the product’s Price category. Statements 1 and 3 are not particularly relevant to the product: Statement 1 is simply a statement of fact, and Statement 3 is a paraphrasing of the product’s slogan. Thus, only Statement 2 provides useful information. The near zero score indicates that there is general disagreement on this statement and the perceptions levels indicate that the players were unsure of how others would perceive the product. This

Table 5: Game 1 ‘Placement’ statements.  $P_i$  represents the response from player  $i$  to the statement in the format of their perception level (1–9) and their own opinion on the truth of the statement. The encoding player responses are in **bold**. The last column represents the statement score.

	Statement	$P_1$	$P_2$	$P_3$	$P_4$	$\Sigma$
1	The product is a snack food	8T	<b>7T</b>	7T	8T	18
2	The product is relatively new on the market	6T	<b>7T</b>	8T	8T	16
3	This product could not be eaten in a rush	6T	<b>7T</b>	8T	8T	16
4	I would still be hungry if I ate this product for lunch	8T	<b>7T</b>	7T	7T	15
5	This would appeal more to women than men	6T	7T	<b>8T</b>	7T	13
6	This product would be eaten at a cocktail party	7T	6F	8T	<b>7T</b>	13
7	I can’t think of an occasion when I would buy this product	7T	7T	5F	<b>7T</b>	9
8	I would be embarrassed to buy this product	6T	7T	4T	<b>4F</b>	2
9	I will buy this product in the future	3F	<b>7F</b>	4F	3F	−4
10	I would feed this product to my children for lunch	5F	6T	<b>3F</b>	3F	−5
11	I have seen this product advertised	4F	4F	<b>2T</b>	3F	−11

Table 6: Game 2 ‘Price’ statements.

	Statement	$P_1$	$P_2$	$P_3$	$P_4$	$\Sigma$
1	The company that makes this product was established many years ago	6T	<b>7T</b>	6T	8T	11
2	The product looks like it’s trying to be expensive	3T	7T	<b>4T</b>	4F	−2
3	I think this product is the happiest drink in the world	<b>1F</b>	4F	4F	7T	−9

indicates that the product does not portray a clear image of how expensive it is.

Game 3 was the first game in Experiment 2. This experiment added a relevance score to each data point, which represented the average perceived relevance of the statement as determined by the encoders. A total of 32 data points was generated. Table 7 contains the statements made in Game 3 relating to the ‘Packaging’ aspects of the product. The high scoring positive statements suggest that the product provides either a mysterious or a confusing image. The mysterious aspect is in line with the branding of the product, however, comments made after the game suggested that the mix between a Latin brand

Table 7: Game 3: ‘Packaging’ statements. The final column represents the perceived relevance of the statement to the product.

	Statement	$P_1$	$P_2$	$P_3$	$P_4$	$\Sigma$	Rel
1	The name is unhelpful – doesn’t say much about the product	7T	6T	<b>7T</b>	7F	10	3.75
2	It isn’t in a large container	8T	<b>6F</b>	3F	7T	7	2.25
3	It looks intriguing	6T	7T	6T	<b>6T</b>	6	3.25
4	The label is attractive	9T	<b>5T</b>	3F	2F	1	4.25
5	The label looks cheap	2F	3F	<b>5T</b>	8T	−3	3.75
6	The [packaging] colour is a bit off-putting	<b>5T</b>	3F	3F	3F	−9	3.75
7	Product doesn’t look very healthy	<b>2F</b>	3F	3F	6F	−12	4.00
8	The bottle is a traditional shape	2F	3F	5T	<b>1F</b>	−19	3.25

name and a Chinese theme were in conflict. Statement 2, a factual statement, was possibly designed to confuse the dummy player. This statement scored very low on relevance, and thus does not provide much information to the design team. The large negative scoring statements provide interesting and positive feedback on the product’s packaging; namely that it is not off-putting, has a healthy appearance, and does not come in a traditionally shaped bottle.

Game 4 provided a total of 32 data points, including relevance scores as per Game 3. Table 8 contains the ‘Product’ related statements made during this game. Most of these statements were encoded as being true, which is contrary to the suggested strategy of mixing true and false statements. However, this game did not dictate when players made statements under particular sections and it is possible the players did not realise they were consistently generating true ‘product’ statements. The majority of the statements were perceived to be relevant, with the exception of Statement 10, which was designed to confuse the dummy player.

Table 8: Game 4: ‘Product’ statements.

	Statement	$P_1$	$P_2$	$P_3$	$P_4$	$\Sigma$	Rel
1	You throw away part of the product after use	<b>8T</b>	9T	9T	9T	36	4.00
2	You would store this product in your home	<b>9T</b>	8T	9T	9T	36	3.50
3	The product is safe to dispose of	8T	<b>6T</b>	7T	9T	20	3.75
4	It has a purposeful appearance	7T	7T	7T	<b>7T</b>	12	3.75
5	It’s bigger than the packaging suggests	7T	7T	<b>5T</b>	7T	9	3.50
6	The product is disappointing compared with the packaging	<b>6T</b>	7T	4F	8T	9	4.00
7	It’s a bit of a gimmick	5F	6T	4F	<b>8T</b>	6	3.50
8	It smells horrible	5F	<b>3F</b>	8T	7F	6	4.00
9	People may find it difficult to use	<b>3F</b>	6T	3F	8T	1	3.75
10	The product is yummy	1F	2F	<b>1F</b>	1F	−36	1.75

## 6.2 Participant game feedback

After the game playing parts of the two experiments were completed, the players were asked to fill in a questionnaire regarding their views on the process. The players were asked to rate their experience according to the following criteria on a scale of 1 (Very Low) to 5 (Very High):

1. Enjoyment of the game;
2. Understanding of the game;
3. Understanding of the scoring;
4. Preference for knowing scores (Experiment 1)/Understanding of relevance of scoring (Experiment 2);
5. Ease of filling out forms; and
6. Helpfulness of the direction supplied.

Criterion 4 related to the score reporting method. During Experiment 1 two approaches were used: in Game 1 the scores were only revealed at the end of the game, while in Game 2 the scores were continuously updated. This was to measure the impact of continuous feedback on the players. In Experiment 2, both games were provided with continuous scoring. Here, the aim was to determine how well the players felt they understood about how their own performance related to the scores they were accumulating.

Criterion 6 related to the prompting supplied by the environment and the moderator. In each game, the encoders were supplied with some additional material to prompt (or inspire) them to make relevant statements about the product. In Experiment 1, this was achieved by displaying a set of keywords to the encoders. In Experiment 2, the statements were directed using the Four P's of marketing (Bayliss; 1985). This experiment tested two approaches for directing the statements: in Game 3 the players were required to make statements using a prespecified category while in Game 4 they were allowed to select which category they were to use for each statement.

Table 9 represents the feedback from Experiment 1. Most of the players enjoyed the game and had a good understanding of the game and the scoring. The mechanical form filling aspect did not represent a problem, which is encouraging as this represents the key to the accessibility of the game. A very strong preference was expressed for being provided with continuous score updates. Clearly, the players felt they benefited and thus maintained an interest in the process when provided continuous updates on their progress. Finally, the players reported that they did not find the direction supplied in this experiment particularly helpful. This is contrary to previous findings (Matthews et al.; 2005), however the earlier work compared the use of keywords to not having any direction at all.

Table 10 represents the feedback from Experiment 2. This group of players had a greater enjoyment and understanding of the game than those who participated in Experiment 1. This appears to be due to the changes made to the game process resulting

Table 9: Post-experiment questionnaire analysis for Experiment 1.

Criterion	Very High	High	Average	Low	Very Low
Enjoyment of the game	3	1		1	
Understanding of the game	1	3		1	
Understanding of scoring		4		1	
Preference for knowing scores	5				
Ease of filling out forms	4	1			
Helpfulness of keywords				4	1

Table 10: Post-experiment questionnaire analysis for Experiment 2.

Criterion	Very High	High	Average	Low	Very Low
Enjoyment of the game	3	2			
Understanding of the game	3	2			
Understanding of scoring	3	1		1	
Understanding relevance of score		4	1		
Ease of filling out forms	5				
Helpfulness of categories		4	1		

from the analysis of Experiment 1. These changes were primarily that the scoring was continuously updated and the introduction and description of the game methodology had been improved. The players preferred the categories to provide them with direction and/or inspiration for generating statements than the keywords approach. Contrary to the hypothesis, players preferred the rigid structure of Game 3 where they were required to make statements within a specific category to the more relaxed approach of Game 4 where they were at liberty to select which category to use. The only player who disagreed with this statement was the ‘last’ in each round, and tended to find that the statements had been ‘used up’ by the time it was their turn.

## 7 Discussion

A significant volume of useful product information was elicited over the course of the two experiments undertaken during this project.

The first objective of Experiment 1 was to examine incentives in the Information Pump in the form of players tracking their scores during the game. From player responses to the post-experiment questionnaire, it was clear that players much preferred knowing their scores during the game. It indicated that a more competitive element had been added, and that this made the game more enjoyable and entertaining. Players' concentration was therefore maintained for longer periods. It also emerged that the one player, who felt she did not understand the game or the scoring system, became quickly disillusioned as her score fell behind the other players.

The second objective of Experiment 1 was to examine the scoring system used in the Information Pump. During Game 1, Prelec's scoring system was used, but players had no knowledge of the manner in which their scores were calculated, nor were they able to track them. This was to prevent the possibility of players calculating that a neutral strategy represented the best average payoff. Players found this situation difficult and frustrating, as they could not develop strategies to improve their score during the game. During Game 2, the Robson (2004) scoring system was put in place. As a neutral strategy no longer represented the highest average payoffs, players were allowed information on how their scores were calculated. This proved far more popular with the majority of players, particularly those who felt they understood the scoring system well.

During Game 2, the Dummy player guessed the identity of the product being discussed. While this did not affect the quality of the information provided, it did result in a very low-scoring game for the Encoders. This was because the Dummy player could accurately predict the other players' responses to the statements posed. As Babycham is a product



with a long history and a recognisable brand, it is not surprising that the Dummy realised the product identity. This indicates that the Information Pump is not suitable for use with well-known products, and is better used for conceptual designs.

Experiment 2 sought to improve upon the flaws discovered during Experiment 1. These included the players finding the keywords unhelpful. In response to this, the categories used in analysis were introduced into the game as a guide for players. Game 3 was split into four defined sections, each relating to one of the analysis categories. While it was predicted that players would find this structured game restrictive in terms of statement composition, the majority of players felt guidance in this form was beneficial. The only player to disagree was the fourth Encoder, who felt that his ideas for statements had been ‘used up’ by the other players. During Game 4, a less structured form of guidance was put in place, with players fitting their own statements into the category they felt was most appropriate. While it was hypothesised that players would find this softer approach more appealing, the opposite was true. Encoders found the extra task of fitting their statement into a category distracting and confusing.

Throughout Experiment 2, players were asked to indicate the relevance of each statement to the product in the defined category, using a scale from 1 to 5. This proved useful in analysis, as the statements that players found to be irrelevant could be identified and excluded. The Information Pump was found to elicit relevant information in each of the four categories, indicating it to be a potentially useful tool in developing the marketing mix of a product. In order for this to be possible, several IP games focusing on the same product concept would be required. For each game, a separate group of participants would be necessary. In this way, it would be possible to elicit a much wider range of customer statements from which the customer needs for the product could be attained. These needs could then be arranged in one of two ways. The first would be to simply group the needs into similar areas (Urban and Hauser; 1993). The second, preferred method would be to

develop the customer needs into, for example, an Objectives Tree as described by (Pahl and Beitz; 1996). Objectives Trees are usually developed by the design team and weightings attached to each of the criterion according to their opinions.

In order to classify the importance of each criterion according to the customer, a measure similar to that used for relevance in Experiment 2 could be added to participant response sheets when playing the IP. This scale would measure the ‘Importance’ of each statement in relation to the product. For example, on a scale of 1 to 5, the value 1 would indicate ‘This aspect would have no influence on my choice of product’, while the value 5 would indicate ‘This aspect would greatly influence my choice of product’. Obtaining quantitative information such as this would be highly beneficial in providing the customer with the product they desire.

It is therefore suggested that the IP be used not only as a method of evaluating products against existing criteria, but also as a method of developing those criteria. In doing so, the IP will become a more powerful tool not only in the fundamental stages of design, but also in marketing research.

## 8 Conclusions

The results obtained suggest that the paper-based version of the Information Pump provides relevant, useful information in a format that is easily accessible.

Experiment 1 proved that the paper-based version of the IP is greatly improved in terms of player enjoyment by displaying scores as the game progresses. Not only does this improve concentration, but also it allows players to adjust their strategies as they see fit. By introducing a measure of statement relevance to the Information Pump in Experiment 2, it was seen that the majority of statements made were considered by the players to be relevant to the product. Statements with a low relevance score could be discarded during analysis,

avoiding distraction from the more useful, relevant statements.

The keywords utilised in Experiment 1 were found to be unhelpful to the players' creativity. These were therefore replaced in Experiment 2 by categories. Players found the structured system of playing the game in rounds relating to these categories preferable to a less structured game. While the categories chosen for this investigation related to the Four Ps of Marketing, categories reflecting the design features of the product may also be used.

It is anticipated that the repeated examination of a single product concept using the Information Pump would be useful in eliciting precise customer needs. These needs could be quantified by introducing an importance scale to the game similar to the relevance scale used during this investigation.

## References

- Bayliss, J. S. (1985). *Marketing for Engineers*, IEE Management of Technology Series, Institution of Electrical Engineers, London.
- Dahan, E. and Hauser, J. R. (2002). The virtual customer, *The Journal of Product Innovation Management* **19**: 332–353.
- Dibb, S., Simkin, L., Pride, W. M. and Ferrell, O. C. (2001). *Marketing: Concepts and Strategies*, Houghton Mifflin.
- Göritz, A. S. (2004). The impact of material incentives on response quantity, response quality, sample composition, survey outcome, and cost in online access panels, *International Journal of Market Research* **46**(3): 327–346.
- Griffin, A. and Hauser, J. R. (1993). The voice of the customer, *Marketing Science* **12**(1): 1–27.

- Hollins, B. and Pugh, S. (1990). *Successful Product Design*, Butterworths.
- Kotler, P. and Armstrong, G. (1997). *Marketing: An Introduction*, fourth edn, Prentice Hall International, Upper Saddle River, NJ.
- Lee, C. K. M., Lau, H. C. W., Yu, K. M. and Ip, W. H. (2005). Enhancement of product development through product information mark-up language, *Journal of Engineering Design* **16**(1): 91–110.
- Lee, E., Hu, M. Y. and Toh, R. S. (2004). Respondent non-cooperation in surveys and diaries: an analysis of item non-response and panel attrition, *International Journal of Market Research* **46**(3): 311–326.
- Matthews, P. C., Keegan, J. D. and Robson, J. R. (2005). Development of a simple information pump, in A. E. Samuel and W. P. Lewis (eds), *Proceedings of the 15th International Conference on Engineering Design*, Design Society, Melbourne.
- Morgan, D. L. (1998). *The Focus Group Guide Book*, Sage, London.
- Pahl, G. and Beitz, W. (1996). *Engineering Design: A Systematic Approach*, second edn, Springer-Verlag London.
- Prelec, D. (2001). A two-person scoring rule for subjective reports, *Technical report*, MIT Sloan School of Management.
- Prelec, D. (2004). A Bayesian truth serum for subjective data, *Science* **206**(5695): 462–466.
- Resnick, P. and Varian, H. (1996). Recommender systems, *Communications of the ACM* **39**: 87–93.
- Robson, J. R. (2004). *Information pump*, Master’s thesis, School of Engineering, University of Durham.

- Romp, G. (1997). *Game Theory: Introduction and Applications*, Oxford University Press, Oxford.
- Siu, K. (2005). Pleasurable products: public space furniture with userfitness, *Journal of Engineering Design* **16**(6): 545–555.
- Spiggle, S. (1994). Analysis and interpretation of qualitative data in consumer research, *Journal of Consumer Research* **21**(3): 491–503.
- Takala, R. (2005). Product demonstrator: a system for up-front testing of user-related product features, *Journal of Engineering Design* **16**(3): 329–336.
- Urban, G. L. and Hauser, J. R. (1993). *Design and Marketing of New Products*, Prentice-Hall International Series in Management, Prentice Hall, London.
- Vehovar, V. and Lozar, K. (1998). How many mailings are enough?, in A. Koch and R. Porst (eds), *Nonresponse in Survey Research*, ZUMA, Mannheim, pp. 139–150.
- Yalch, R. and Brunel, F. (1996). Need hierarchies in consumer judgments of product designs: Is it time to reconsider Maslow’s theory?, *Advances in Consumer Research* **23**: 405–410.
- Yates, J. F. (ed.) (1992). *Risk-taking behavior*, Wiley series in Human Performance and Cognition, Wiley, Chichester, NY.
- Zikmund, W. G. (2002). *Essentials of Marketing Research*, The Dryden Press Series in Marketing, South Western College Publishing.